(72) Inventors:
• Kushida, Akihiro, c/o Canon Kabushiki Kaisha
Ohta-ku, Tokyo (JP)
• Kosaka, Tetsuo, c/o Canon Kabushiki Kaisha
Ohta-ku, Tokyo (JP)

(74) Representative:
Beresford, Keith Denis Lewis et al
BERESFORD & Co,
2-5 Warwick Court,
High Holborn
London WC1R 5DH (GB)

(54) **Client-server based speech recognition**

(57) A user dictionary, which is formed by storing pronunciations and notations of target recognition words designated by the user in correspondence with each other, input speech recognition data, and dictionary management data used to determine the recognition field of a recognition dictionary used in recognition of the speech recognition data are sent to a server via a communication module. In the server, a dictionary management unit looks up an identifier table to determine a recognition dictionary corresponding to the dictionary management information received from a client from a plurality of kinds of recognition dictionaries. A speech recognition module recognizes the speech recognition data using at least the determined recognition dictionary. The recognition result is sent to the client via a communication module.

FIG. 6
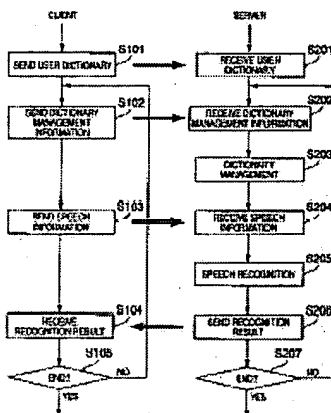
EP 1 209 662 A2

Description

FIELD OF THE INVENTION

[0001] The present invention relates to a client-server speech recognition system for recognizing speech input at a client by a server, a speech recognition server, a speech recognition client, their control method, and a computer readable memory.

BACKGROUND OF THE INVENTION

[0002] In recent years, speech is used as an input interface in addition to a keyboard, mouse, and the like.

[0003] However, the recognition rate of speech recognition that recognizes input speech lowers and requires a longer processing time as the number of recognition words which are to undergo speech recognition becomes larger. For this reason, in an actual method, a plurality of recognition dictionaries or lexicons that register recognition words (e.g., pronunciations and notations) which are to undergo speech recognition are prepared, and are selectively used (a plurality of recognition dictionaries may be used at the same time).

[0004] Also, unregistered words cannot be recognized. As one of methods for solving this problem, a user dictionary or lexicon (prepared by the user to register recognition words which are to undergo speech recognition) may be used.

[0005] On the other hand, a client-server speech recognition system has been studied to implement speech recognition on a terminal with insufficient resources.

[0006] These three techniques are known to those who are skilled in the art, but a system that combines these three techniques has not been realized yet.

SUMMARY OF THE INVENTION

[0007] The present invention has been made to solve the above problems, and has as its object to provide a speech recognition system which uses a user dictionary in response to a user's request in a client-server speech recognition system so as to improve speech input efficiency and to reduce the processing load on the entire system, a speech recognition server, a speech recognition client, their control method, and a computer readable memory.

[0008] According to the present invention, the foregoing object is attained by providing, a client-server speech recognition system for recognizing speech input at a client by a server,

the client comprising:

speech input means for inputting speech;
user dictionary holding means for holding a user dictionary formed by registering target recognition words designated by a user; and

transmission means for transmitting speech data input by said speech input means, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary to the server, and

the server comprising:

recognition dictionary holding means for holding a plurality of kinds of recognition dictionaries prepared for respective recognition fields;
determination means for determining one or more recognition dictionary corresponding to the dictionary management information received from the client from the plurality of kinds of recognition dictionaries; and
recognition means for recognizing the speech data using at least the recognition dictionary determined by said determination means.

[0009] Other features and advantages of the present invention will be apparent from the following description taken in conjunction with the accompanying drawings, in which like reference characters designate the same or similar parts throughout the figures thereof.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010]

Fig. 1 is a block diagram showing the hardware arrangement of a speech recognition system of the first embodiment;
Fig. 2 is a block diagram showing the functional arrangement of the speech recognition system of the first embodiment;
Fig. 3 shows the configuration of a user dictionary of the first embodiment;
Fig. 4 shows a speech input window of the first embodiment;
Fig. 5 shows an identifier table of the first embodiment;
Fig. 6 is a flow chart showing the process executed by the speech recognition system of the first embodiment;
Fig. 7 shows the configuration of a user dictionary appended with input form identifiers according to the third embodiment; and
Fig. 8 shows the configuration of a user dictionary appended with recognition dictionary identifiers according to the third embodiment.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0011] Preferred embodiments of the present invention will be described in detail below with reference to

the accompanying drawings.

[First Embodiment]

[0012] Fig. 1 shows the hardware arrangement of a speech recognition system of the first embodiment.

[0013] A CPU 101 systematically controls an entire client 100. The CPU 101 loads programs stored in a ROM 102 onto a RAM 103, and executes various processes on the basis of the loaded programs. The ROM 102 stores various programs of processes to be executed by the CPU 101. The RAM 103 provides a storage area required to execute various programs stored in the ROM 102.

[0014] A secondary storage device 104 stores an OS and various programs. When the client 100 is implemented using not a general-purpose apparatus such as a personal computer or the like but a dedicated apparatus, the ROM 102 may store the OS and various programs. By loading the stored programs onto the RAM 103, the CPU 101 can execute processes. As the secondary storage device 104, a hard disk device, floppy disk drive, CD-ROM, or the like may be used. That is, storage media are not particularly limited.

[0015] A network I/F (interface) 105 is connected to a network I/F 205 of a server 200.

[0016] An input device 106 comprises a mouse, keyboard, microphone, and the like to allow input of various instructions to processes to be executed by the CPU 101, and can be used by simultaneously connecting these plurality of devices. An output device 107 comprises a display (CRT, LCD, or the like), and displays information input by the input device 106, and display windows which are controlled by various processes executed by the CPU 101. A bus 108 interconnects various building components of the client 100.

[0017] A CPU 201 systematically controls the entire server 200. The CPU 201 loads programs stored in a ROM 202 onto a RAM 203, and executes various processes on the basis of the loaded programs. The ROM 202 stores various programs of processes to be executed by the CPU 201. The RAM 203 provides a storage area required to execute various programs stored in the ROM 202.

[0018] A secondary storage device 204 stores an OS and various programs. When the server 200 is implemented using not a versatile apparatus such as a personal computer or the like but a dedicated apparatus, the ROM 202 may store the OS and various programs. By loading the stored programs onto the RAM 203, the CPU 201 can execute processes. As the secondary storage device 204, a hard disk device, floppy disk drive, CD-ROM, or the like may be used. That is, storage media are not particularly limited.

[0019] The network I/F 205 is connected to the network I/F 105 of the client 100. A bus 206 interconnects various building components of the server 200.

[0020] The functional arrangement of the speech rec-

ognition system of the first embodiment will be described below using Fig. 2.

[0021] Fig. 2 is a block diagram showing the functional arrangement of the speech recognition system of the first embodiment.

[0022] In the client 100, a speech input module 121 inputs speech uttered by the user via a microphone (input device 106), and A/D-converts input speech data (speech recognition data) which is to undergo speech recognition. A communication module 122 sends a user dictionary 124a, speech recognition data 124b, dictionary management information 124c, and the like to the server 200. Also, the communication module 122 receives a speech recognition result of the sent speech recognition data 124b and the like from the server 200.

[0023] A display module 123 displays the speech recognition result received from the server 200 while storing it in, e.g., an input form which is displayed on the output device 107 by the process executed by the speech recognition system of this embodiment.

[0024] In the server 200, a communication module 221 receives the user dictionary 124a, speech recognition data 124b, dictionary management information 124c, and the like from the client 100. Also, the communication module 221 sends the speech recognition result of the speech recognition data 124b and the like to the client 100.

[0025] A dictionary management module 223 switches and selects a plurality of kinds of recognition dictionaries 225 (recognition dictionary 1 to recognition dictionary N, N: a positive integer) prepared for respective recognition fields (e.g., for names, addresses, alphanumeric symbols, and the like), and the user dictionary 124a received from the client 100 (may simultaneously use a plurality of kinds of dictionaries).

[0026] Note that the plurality of kinds of recognition dictionaries 225 are prepared for each dictionary management information 124c (input form identifier; to be described later) sent from the client 100. Each recognition dictionary 225 is appended with a recognition dictionary identifier indicating the recognition field of that recognition dictionary. The dictionary management module 223 manages an identifier table 223a that stores the recognition dictionary identifiers and input form identifiers in correspondence with each other, as shown in Fig. 5.

[0027] A speech recognition module 224 executes speech recognition using the recognition dictionary or dictionaries 225 and user dictionary 124a designated for speech recognition by the dictionary management module 223 on the basis of the speech recognition data 124b and dictionary management information 124c received from the client 100.

[0028] Note that the user dictionary 124a is prepared by the user to register recognition words which are to undergo speech recognition, and stores pronunciations and notations of words to be recognized in correspondence with each other, as shown in, e.g., Fig. 3.

[0029] The speech recognition data 124b may be either speech data A/D-converted by the speech input module 121 or data obtained by encoding that speech data.

[0030] The dictionary management information 124c indicates an input object and the like. For example, the dictionary management information 124c is an identifier (input form identifier) indicating the type of input form when the server 200 recognizes input speech and inputs text data corresponding to that speech recognition result to each input form, which defines a speech input window displayed by the speech recognition system of the first embodiment, as shown in Fig. 4. The client 100 sends this input form identifier to the server 200 as the dictionary management information 124c. In the server 200, the dictionary management module 223 looks up the identifier table 223a to acquire a recognition dictionary identifier corresponding to the received input form identifier, and determines a recognition dictionary 225 to be used in speech recognition.

[0031] The process executed by the speech recognition system of the first embodiment will be explained below using Fig. 6.

[0032] Fig. 6 is a flow chart showing the process executed by the speech recognition system of the first embodiment.

[0033] In step S101, the client 100 sends the user dictionary 124a to the server 200.

[0034] In step S201, the server 200 receives the user dictionary 124a from the client 100.

[0035] In step S102, when speech is input to an input form as a target speech input, the client 100 sends the input form identifier of that input form to the server 200 as the dictionary management information 124c.

[0036] In step S202, the server 200 receives the input form identifier from the client 100 as the dictionary management information 124c.

[0037] In step S203, the server 200 looks up the identifier table 223a using the dictionary management information 124c to acquire a recognition dictionary identifier corresponding to the received input form identifier, and determines a recognition dictionary 225 to be used in speech recognition.

[0038] In step S103, the client 100 sends speech recognition data 124b, which is speech-input as text data to be input to each input form, to the server 200.

[0039] In step S204, the server 200 receives the speech recognition data corresponding to each input form from the client 100.

[0040] In step S205, the server 200 executes speech recognition of the speech recognition data 124b in the speech recognition module 224 using the recognition dictionary 225 and user dictionary 124a designated for speech recognition by the dictionary management module 223.

[0041] In the first embodiment, all recognition words contained in the user dictionary 124a sent from the client 100 to the server 200 are used in speech recognition by

the speech recognition module 224.

[0042] In step S206, the server 200 sends the speech recognition result obtained by the speech recognition module 224 to the client 100.

[0043] In step S104, the client 100 receives the speech recognition result corresponding to each input form from the server 200, and stores text data corresponding to the speech recognition result in the corresponding input form.

[0044] The client 100 checks in step S105 if the processing is to be ended. If the processing is not to be ended (NO in step S105), the flow returns to step S102 to repeat the processing. On the other hand, if the processing is to be ended (YES in step S105), the client 100 informs the server 200 of end of the processing, and ends the processing.

[0045] It is checked in step S207 if a processing end instruction from the client 100 is detected. If no processing end instruction is detected (NO in step S207), the flow returns to step S202 to repeat the above processes. On the other hand, if the processing end instruction is detected (YES in step S207), the processing ends.

[0046] In the above processing, when speech is input to an input form as a target speech input, the dictionary management information 124c corresponding to that input form is sent from the client 100 to the server 200. Alternatively, the dictionary management information 124c may be sent when the input form as a target speech input is focused by an instruction from the input device 106 (the input form as a target speech input is determined).

[0047] In the server 200, speech recognition is made after all speech recognition data 124b are received. Alternatively, every time speech is input as text data to a given input form, that the portion of speech recognition data 124b may be sent to the server 200 frame by frame (for example, one frame is 10 msec speech data), and speech recognition may be made in real time.

[0048] As described above, according to the first embodiment, in the client-server speech recognition system, since the server 200 executes speech recognition of speech recognition data 124b using both an appropriate recognition dictionary 225 and the user dictionary 124a, the speech recognition precision in the server 200 can be improved while reducing the processing load and use of storage resources associated with speech recognition in the client 100.

[Second Embodiment]

[0049] In the first embodiment, if no recognition words to be stored in the user dictionary 124a are generated, since the user dictionary 124a need not be used, the server 200 may use all recognition words in the user dictionary 124a in recognition only when a use request of the user dictionary 124a is received from the client 100.

[0050] In this case, a flag indicating if the user dictionary 124a is used is added as the dictionary management

Information 124c, thus informing the server 200 of the presence/absence of use of the user dictionary 124a.

[Third Embodiment]

[0051] Since some target recognition words in the user dictionary 124a are not used depending on an input object, situation, and the like, only specific recognition words in the user dictionary 124a may be used in recognition depending on the input object and situation.
[0052] In such case, when the user dictionary is managed by designating input form identifiers for respective recognition words, as shown in Fig. 7, only recognition words having an input form identifier of the input form used in speech input can be used in recognition. Alternatively, a plurality of input form identifiers may be designated for a given recognition word. In addition, the user dictionary may be managed by designating recognition dictionary identifiers in place of input form identifiers, as shown in Fig. 8.

[Fourth Embodiment]

[0053] By combining the second and third embodiments, the efficiency of the speech recognition process of the speech recognition module 224 can be further improved.

[Fifth Embodiment]

[0054] Most of the processes of the apparatus of the present invention can be implemented by programs. As described above, since the apparatus can use a general-purpose apparatus such as a personal computer, the present invention is also achieved by supplying a storage medium, which records a program code of a software program that can implement the functions of the above-mentioned embodiments to a system or apparatus, and reading out and executing the program code stored in the storage medium by a computer of the system or apparatus. In this case, the program code itself read out from the storage medium implements the functions of the above-mentioned embodiments, and the storage medium which stores the program code constitutes the present invention. As the storage medium for supplying the program code, for example, a floppy disk, hard disk, optical disk, magneto-optical disk, CD-ROM, magnetic tape, nonvolatile memory card, ROM, and the like may be used.
[0055] The present invention can also be achieved by supplying the storage medium that records the program code to a computer, and executing some or all of actual processes executed by an OS running on the computer. Furthermore, the functions of the above-mentioned embodiments may be implemented by some or all of actual processing operations executed by a CPU or the like arranged in a function extension board or a function extension unit, which is inserted in or connected to the

computer, after the program code read out from the storage medium is written in a memory of the extension board or unit. When the present invention is applied to the storage medium, that storage medium stores a program code corresponding to the flow chart shown in Fig. 3.
[0056] As many apparently widely different embodiments of the present invention can be made without departing from the spirit and scope thereof, it is to be understood that the invention is not limited to the specific embodiments thereof except as defined in the appended claims.

Claims

1. A client-server speech recognition system for recognizing speech input at a client by a server,

   the client comprising:

      speech input means for inputting speech;
      user dictionary holding means for holding a user dictionary formed by registering target recognition words designated by a user; and
      transmission means for transmitting speech data input by said speech input means, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary to the server, and

   the server comprising:

      recognition dictionary holding means for holding a plurality of kinds of recognition dictionaries prepared for respective recognition fields;
      determination means for determining one or more recognition dictionary corresponding to the dictionary management information received from the client from the plurality of kinds of recognition dictionaries; and
      recognition means for recognizing the speech data using at least the recognition dictionary determined by said determination means.

2. The system according to claim 1, wherein said recognition means recognizes the speech data using the recognition dictionary determined by said determination means, and the user dictionary received from the client.

3. The system according to claim 1 or 2, wherein said

speech input means comprises display means for displaying an input form as a target speech input, and

the dictionary management information is an input form identifier that indicates a type of input form.

4. The system according to any of claims 1 to 3, wherein the dictionary management information contains information indicating if the user dictionary is used in recognition of the speech data.

5. The system according to any preceding claim, wherein the user dictionary is formed by storing pronunciations and notations of the target recognition words in correspondence with each other.

6. The system according to claim 3, wherein the user dictionary is formed by also storing at least one input form identifier and the target recognition words in correspondence with each other.

7. The system according to any preceding claim, wherein the user dictionary is formed by also storing at least one of recognition dictionary identifiers indicating recognition fields of the plurality of kinds of recognition dictionaries, and the target recognition words.

8. The system according to any preceding claim, wherein the speech data is data obtained by encoding that speech data.

9. A method of controlling a client-server speech recognition system for recognizing speech input at a client by a server, comprising:

a speech input step of inputting speech;
a user dictionary holding step of holding, in the client, a user dictionary formed by registering target recognition words designated by a user; and
a transmission step of transmitting speech data input in the speech input step, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary to the server;
a recognition dictionary holding step of holding, in the server, a plurality of kinds of recognition dictionaries prepared for respective recognition fields;
a determination step of determining one or more recognition dictionary corresponding to the dictionary management information received from the client from the plurality of kinds of recognition dictionaries; and
a recognition step of recognizing the speech

data using at least the recognition dictionary determined in the determination step.

10. The method according to claim 9, wherein the recognition step includes a step of recognizing the speech data using the recognition dictionary determined in the determination step, and the user dictionary received from the client.

11. The method according to claim 9 or 10, wherein the speech input step comprises a display step of displaying an input form as a target speech input, and
the dictionary management information is an input form identifier that indicates a type of input form.

12. The method according to any of claims 9 to 11, wherein the dictionary management information contains information indicating if the user dictionary is used in recognition of the speech data.

13. The method according to any of claims 9 to 12, wherein the user dictionary is formed by storing pronunciations and notations of the target recognition words in correspondence with each other.

14. The method according to claim 11, wherein the user dictionary is formed by also storing at least one input form identifier and the target recognition words in correspondence with each other.

15. The method according to any of claims 9 to 14, wherein the user dictionary is formed by also storing at least one of recognition dictionary identifiers indicating recognition fields of the plurality of kinds of recognition dictionaries, and the target recognition words.

16. The method according to any of claims 9 to 15, wherein the speech data is data obtained by encoding that speech data.

17. A computer readable memory that stores a program code of control of a client-server speech recognition system for recognizing speech input at a client by a server, comprising:

a program code of a speech input step of inputting speech;
a program code of a user dictionary holding step of holding, in the client, a user dictionary formed by registering target recognition words designated by a user; and
a program code of a transmission step of transmitting speech data input in the speech input step, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data,

and the user dictionary to the server;
a program code of a recognition dictionary holding step of holding, in the server, a plurality of kinds of recognition dictionaries prepared for respective recognition fields;
a program code of a determination step of determining one or more recognition dictionary corresponding to the dictionary management information received from the client from the plurality of kinds of recognition dictionaries; and
a program code of a recognition step of recognizing the speech data using at least the recognition dictionary determined in the determination step.

18. A speech recognition server for recognizing speech input at a client, and sending a recognition result to the client, comprising:

reception means for receiving, from the client, speech data, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary formed by registering target recognition words designated by a user;
recognition dictionary holding means for holding a plurality of kinds of recognition dictionaries prepared for respective recognition fields;
determination means for determining one or more recognition dictionary corresponding to the dictionary management information received from the client from the plurality of kinds of recognition dictionaries; and
recognition means for recognizing the speech data using at least the recognition dictionary determined by said determination means.

19. The server according to claim 18, wherein said recognition means recognizes the speech data using the recognition dictionary determined by said determination means, and the user dictionary received from the client.

20. The server according to claim 18 or 19, wherein the speech data is data obtained by encoding that speech data.

21. A speech recognition client for sending input speech to be recognized to a server, and receiving a recognition result of that speech, comprising:

speech input means for inputting speech;
user dictionary holding means for holding a user dictionary formed by registering target recognition words designated by a user; and
transmission means for transmitting speech data input by said speech input means, diction-

ary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary to the server.

22. The client according to claim 21, wherein said speech input means comprises display means for displaying an input form as a target speech input, and
the dictionary management information is an input form identifier that indicates a type of input form.

23. The client according to claim 21 or 22, wherein the dictionary management information contains information indicating if the user dictionary is used in recognition of the speech data.

24. The client according to any of claims 21 to 23, wherein the user dictionary is formed by storing pronunciations and notations of the target recognition words in correspondence with each other.

25. The client according to claim 22, wherein the user dictionary is formed by also storing at least one input form identifier and the target recognition words in correspondence with each other.

26. The client according to any of claims 21 to 25, wherein the user dictionary is formed by also storing at least one of recognition dictionary identifiers indicating recognition fields of the plurality of kinds of recognition dictionaries, and the target recognition words.

27. The client according to any of claims 21 to 25, wherein the speech data is data obtained by encoding that speech data.

28. A method of controlling a speech recognition server for recognizing speech input at a client, and sending a recognition result to the client, comprising:

a reception step of receiving, from the client, speech data, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary formed by registering target recognition words designated by a user;
a recognition dictionary holding step of holding a plurality of kinds of recognition dictionaries prepared for respective recognition fields;
a determination step of determining one or more recognition dictionary corresponding to the dictionary management information received from the client from the plurality of kinds of recognition dictionaries; and

a recognition step of recognizing the speech data using at least the recognition dictionary determined in the determination step.

29. The method according to claim 28, wherein the recognition step includes a step of recognizing the speech data using the recognition dictionary determined in the determination step, and the user dictionary received from the client.

30. The method according to claim 28 or 29, wherein the speech data is data obtained by encoding that speech data.

31. A method of controlling a speech recognition client for sending input speech to be recognized to a server, and receiving a recognition result of that speech, comprising:

a speech input step of inputting speech;
a user dictionary holding step of holding a user dictionary formed by registering target recognition words designated by a user; and
a transmission step of transmitting speech data input in the speech input step, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary to the server.

32. The method according to claim 31, wherein the speech input step comprises a display step of displaying an input form as a target speech input, and the dictionary management information is an input form identifier that indicates a type of input form.

33. The method according to claim 31 or 32, wherein the dictionary management information contains information indicating if the user dictionary is used in recognition of the speech data.

34. The method according to any of claims 31 to 33, wherein the user dictionary is formed by storing pronunciations and notations of the target recognition words in correspondence with each other.

35. The method according to claim 32, wherein the user dictionary is formed by also storing at least one input form identifier and the target recognition words in correspondence with each other.

36. The method according to any of claims 31 to 35, wherein the user dictionary is formed by also storing at least one of recognition dictionary identifiers indicating recognition fields of the plurality of kinds of recognition dictionaries, and the target recognition words.

37. The method according to any of claims 31 to 36, wherein the speech data is data obtained by encoding that speech data.

38. A computer readable memory that stores a program code of control of a speech recognition server for recognizing speech input at a client, and sending a recognition result to the client, comprising:

a program code of a reception step of receiving, from the client, speech data, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary formed by registering target recognition words designated by a user;
a program code of a recognition dictionary holding step of holding a plurality of kinds of recognition dictionaries prepared for respective recognition fields;
a program code of a determination step of determining one or more recognition dictionary corresponding to the dictionary management information received from the client from the plurality of kinds of recognition dictionaries; and
a program code of a recognition step of recognizing the speech data using at least the recognition dictionary determined in the determination step.

39. A computer readable memory that stores a program code of control of a speech recognition client for sending input speech to be recognized to a server, and receiving a recognition result of that speech, comprising:

a program code of a speech input step of inputting speech;
a program code of a user dictionary holding step of holding a user dictionary formed by registering target recognition words designated by a user; and
a program code of a transmission step of transmitting speech data input in the speech input step, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary to the server.

40. A client-server speech recognition system for recognizing speech input at a client by a server,

the client comprising:

a speech input unit inputs speech;
a user dictionary holding a user dictionary formed by registering target recognition words designated by a user; and

a transmitter transmits speech data input by said speech input means, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary to the server, and

the server comprising:

a recognition dictionary holding a plurality of kinds of recognition dictionaries prepared for respective recognition fields;
a determination unit determines one or more recognition dictionary corresponding to the dictionary management information received from the client from the plurality of kinds of recognition dictionaries; and
a recognition unit recognizes the speech data using at least the recognition dictionary determined by said determination means.

41. A speech recognition server for recognizing speech input at a client, and sending a recognition result to the client, comprising:

a receiver receives, from the client, speech data, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and a user dictionary formed by registering target recognition words designated by a user;
a recognition dictionary holding a plurality of kinds of recognition dictionaries prepared for respective recognition fields;
a determination unit determines one or more recognition dictionary corresponding to the dictionary management information received from the client from the plurality of kinds of recognition dictionaries; and
a recognition unit recognizes the speech data using at least the recognition dictionary determined by said determination means.

42. A speech recognition client for sending input speech to be recognized to a server, and receiving a recognition result of that speech, comprising:

a speech input unit inputs speech;
a user dictionary holding a user dictionary formed by registering target recognition words designated by a user; and
a transmitter transmits speech data input by said speech input means, dictionary management information used to determine a recognition field of a recognition dictionary used to recognize the speech data, and the user dictionary to the server.
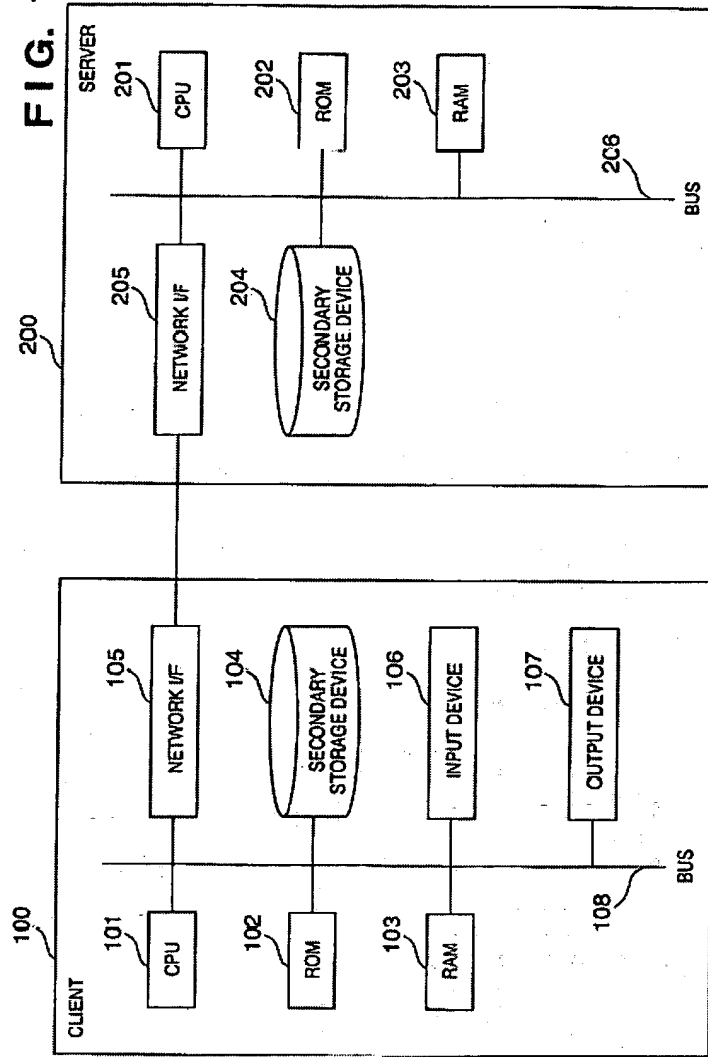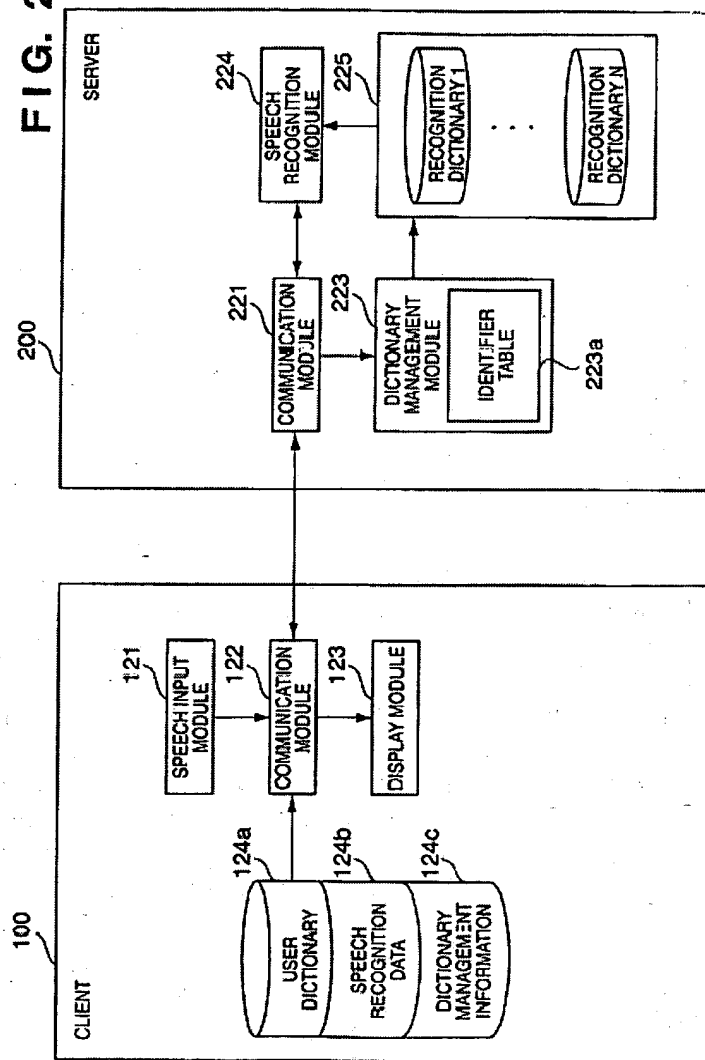
43. Processor implementable instructions product for causing a programmable computer device to carry out the method of any of claims 28 to 37, when the instructions product is run on said programmable computer device.

FIG. 1

SERVER 200

- 201 CPU
- 202 ROM
- 203 RAM
- 205 NETWORK I/F
- 204 SECONDARY STORAGE DEVICE
- 206 BUS

CLIENT 100

- 101 CPU
- 102 ROM
- 103 RAM
- 105 NETWORK I/F
- 104 SECONDARY STORAGE DEVICE
- 106 INPUT DEVICE
- 107 OUTPUT DEVICE
- 108 BUS

FIG. 2

SERVER 200

COMMUNICATION MODULE 221

SPEECH RECOGNITION MODULE 224

DICTIONARY MANAGEMENT MODULE 223

IDENTIFIER TABLE 223a

RECOGNITION DICTIONARY 1 225

RECOGNITION DICTIONARY N

. . .

CLIENT 100

SPEECH INPUT MODULE 121

COMMUNICATION MODULE 122

DISPLAY MODULE 123

USER DICTIONARY 124a

SPEECH RECOGNITION DATA 124b

DICTIONARY MANAGEMENT INFORMATION 124c

FIG. 3

USER DICTIONARY

PRONUNCIATION	NOTATION
スズキタロー	鈴木太郎
シタクジュウショ	神奈川県川崎市中原区今井上町999
シタクデンワ	045-222-2222
ケイタイデンワ	090-1111-1111
ファックス	045-333-3333
メール	suzuki@xxx.yyy.co.jp

FIG. 4

FIG. 5

INPUT FORM IDENTIFIER	RECOGNITION DICTIONARY IDENTIFIER
NAME	NAME
ADDRESS	ADDRESS
TEL	ALPHANUMERIC SYMBOL
FAX	ALPHANUMERIC SYMBOL
E-MAIL	ALPHANUMERIC SYMBOL
OTHER	DICTATION
	: :

FIG. 6

CLIENT   SERVER

SEND USER DICTIONARY — S101	RECEIVE USER DICTIONARY — S201
SEND DICTIONARY MANAGEMENT INFORMATION — S102	RECEIVE DICTIONARY MANAGEMENT INFORMATION — S202
	DICTIONARY MANAGEMENT — S203
SEND SPEECH INFORMATION — S103	RECEIVE SPEECH INFORMATION — S204
	SPEECH RECOGNITION — S205
RECEIVE RECOGNITION RESULT — S104	SEND RECOGNITION RESULT — S206
END? S105 NO	END? S207 NO
YES	YES
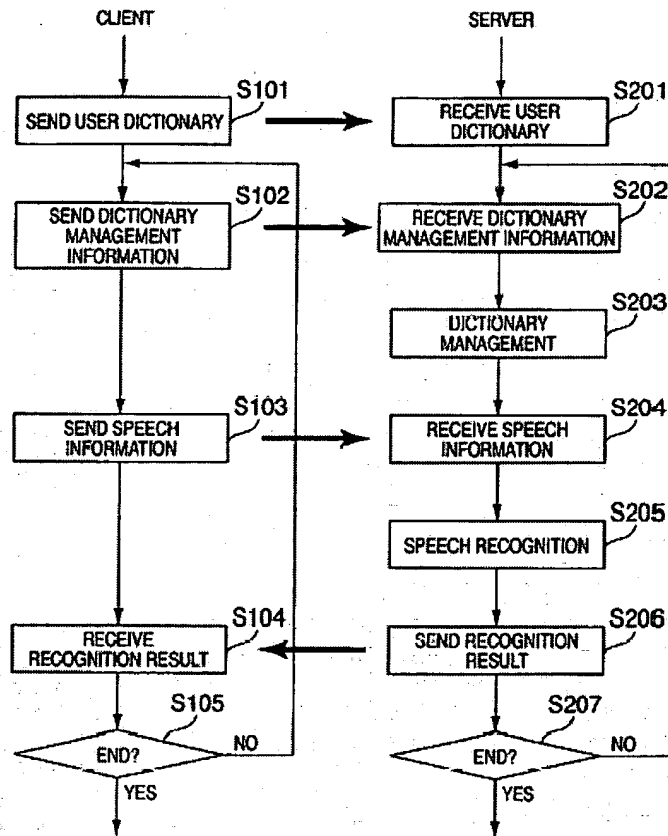
F I G. 7

USER DICTIONARY

PRONUNCIATION	NOTATION	INPUT FORM IDENTIFIER
スズキタロー	鈴木太郎	NAME
ジタククシュウショ	神奈川県川崎市中原区今井上町999	ADDRESS
ジタクデンワ	045-222-2222	TEL
ケイタイデンワ	090-1111-1111	TEL
ファックス	045-333-3333	FAX
メール	suzuki@xxx.yyy.co.jp	E-MAIL
スズキタロー	鈴木太郎	OTHER

FIG. 8

USER DICTIONARY

PRONUNCIATION	NOTATION	RECOGNITION DICTIONARY IDENTIFIER
スズキタロー	鈴木太郎	NAME
ジタクジュウショ	神奈川県川崎市中原区今井上町53	ADDRESS
ジタクデンワ	045-222-2222	ALPHANUMERIC SYMBOL
ケイタイデンワ	090-1111-1111	ALPHANUMERIC SYMBOL
ファックス	045-333-3333	ALPHANUMERIC SYMBOL
メール	suzuki@xxx.yyy.oo.jp	ALPHANUMERIC SYMBOL
スズキタロー	鈴木太郎	DICTATION